A STUDY OF THE RELATIVE EFFECTS OF AGE AND OF TEST DIFFICULTY UPON FACTOR PATTERNS

By

HAZEN ALONZO CURTIS

B.Ed., Southern Illinois Normal University, 1934
M.S., University of Illinois, 1938
Ed.M., University of Illinois, 1947

A SUMMARY OF A THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF EDUCATION
IN THE GRADUATE COLLEGE OF THE
UNIVERSITY OF ILLINOIS, 1948

Reprinted from *Genetic Psychology Monographs*
Vol. 40, No. 1, pages 101-148, 1949

Published as a separate and in *Genetic Psychology Monographs,* 1949, **40**, 99-148.

A STUDY OF THE RELATIVE EFFECTS OF AGE AND OF TEST DIFFICULTY UPON FACTOR PATTERNS*

School of Education, University of Illinois

HAZEN ALONZO CURTIS

Acknowledgments 101

I. Setting of the problem 103

II. The experiment 113

III. Description of the data 123

IV. The factorial analyses 131

V. Interpretation and conclusions 143

References 147

*Received in the Editorial Office on February 1, 1949. Copyright, 1949, by The Journal Press, Provincetown, Massachusetts.

[1]Submitted in partial fulfillment of the requirements for the degree of Doctor of Education in the Graduate College of the University of Illinois, 1948.

R. UNIV.
BIBLIOTHEEK
LEIDEN

ACKNOWLEDGMENTS

The investigator wishes to acknowledge his indebtedness to Professor Edwin H. Reeder, Chairman of his Committee, and to Professors F. H. Finch, G. M. Blair, and E. L. Welker, Members, who by their initial interest and continued guidance made this study possible. It is also a pleasure to acknowledge the contribution of the Administrators and Teachers of the Public Schools of Staunton and of Springfield, Illinois, and of their pupils who served as subjects. Many others contributed their services and advice, both of which have added much to the value of this study.

HAZEN ALONZO CURTIS

School of Education
The Florida State University
Tallahassee, Florida

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41549807_0089

I. THE SETTING OF THE PROBLEM

A. The Hypotheses

Psychologists over the past 20 years have studied the relationship between the age of subjects and the factor patterns obtained from their scores on batteries of intelligence tests. Professor Henry E. Garrett, in his address to the American Psychological Association as its retiring president for the year 1946, presented evidence to support a differentiation hypothesis with respect to the growth of intelligence. Professor Garrett's hypothesis was: "Abstract or symbol intelligence changes in its organization as age increases from a fairly unified and general ability to a loosely organized group of abilities or factors" (12, p. 373). From his hypothesis Professor Garrett drew important theoretical as well as practical implications. Briefly, he pointed out that it "seems to effect a rapprochment between the Spearman General Factor and the Group Factor theories," and further that "Thorndike's 'quantity hypothesis' is not opposed to the differentiation hypothesis. . . ." (12, p. 376).

Concerning its implications for practice he said:

> "The differentiation hypothesis has definite practical implication for the interpretation of intelligence test scores over a wide age range. . . . I do not think that at the elementary school level we should attempt, except very tentatively, to fractionate the *IQ* into say, language ability, number ability, reasoning and the like" (12, p. 376). . . . At the high school and college levels intelligence breaks down . . . into a number of relatively independent factors. It would seem to be theoretically more defensible, therefore, and practically more useful, to measure verbal, numerical, perceptual, and spatial ability, and perhaps other factors at these ages, than to give the subject a single over-all score" (12, p. 377).

To this investigator it seemed that the implications of this hypothesis were of such importance for both theory and practice that an independent test of it was justified.

In addition, an examination of the testing techniques employed in the studies upon which the differentiation hypothesis was based led the investigator to formulate a second hypothesis to be tested at the same time. This second hypothesis was: The observed differences between the factor patterns for two age groups are due not only to the age differences of the groups in question but also to the differences in the difficulty levels of the test materials for the two populations.

B. Previous Work on the Differentiation Hypothesis

A survey of that part of the literature which concerns the change of mental pattern with age might well begin with Slocombe (31) who was one of the first to consider this problem. Slocombe studied the variations in "*g*" as measured by tests administered to subjects at intervals of time varying from one day to three months. Although we today would look upon this method as being entirely inadequate, Slocombe concluded ". . . that '*g*' really is constant, and that any apparent variations found by testing are attributed to varying group factors involved in the tests used, and that the appearance of such variation is to be regarded as a serious defect in the tests" (31, p. 110).

As long as changes in group factors were considered defects in tests, little in the way of constructive work on the problem of their relationship with the general factor could be expected. It was not long, however, before psychologists became interested in the importance of group factors as indicators of special abilities over and above the general ability designated as "*g*."

T. L. Kelley (19) was one of the first to study their importance as measures of intelligence. He administered a battery of 10 tests to a population of 140 seventh grade subjects and also to 110 third grade subjects of both sexes. Because only 60 of the third graders turned in usable papers on the tenth test, Kelley made two separate analyses of his data. From these he concluded, ". . . the factors found are independent of the tutelage occurring between grade three and grade seven. . . . It thus seems that factors reported are established early in life. In fact it seems reasonable to attribute them, at least in major portion, to original nature" (19, pp. 133-134). When he extended his study to a kindergarten population he again concluded that all of the factors he found in the younger subjects, except one, were of the same nature as those revealed in the older populations (19, p. 149).

Soon after the work of Kelley others became interested in the organization of general and group factors. Schneck (30), in working with college men, found a correlation of only .2625 between verbal and non-verbal abilities, and even lower correlations between these and memory. He concluded that at the age level of his subjects (18 to 21 years) the verbal and number tests were not measures of the same general ability, and thereby indicated that there was a rather complete independence of mental factors in adults.

In contrast to these findings, Bryan (5) reported that at the age of five years intercorrelations between memory tests, vocabulary tests, and the

Stanford Binet yielded a single central factor. The fact that the Stanford Binet and the vocabulary tests showed a slightly higher correlation with this factor than did the memory tests is not so important here as is the fact that these three types of tests all measured much the same thing in young children.

Schiller (29) tested boys and girls aged nine years and found definite evidence of verbal, number, and spatial abilities with ". . . the possibility of a single common function throughout all such tests" (29, p. 58).

Three years later Asch (1) retested 79 of the boys and 82 of the girls of Schiller's original group. He found that a drop in the mean intercorrelations of the tests had taken place, and that this drop was somewhat larger for boys than for girls. He also analyzed the intercorrelations using the multiple factor method, and the results of this analysis fully confirmed both of the previously mentioned facts. These findings are extremely important because Asch's study is one of the few that is based upon a retest of the same population. It should be pointed out here that his results, showing that the decrease between the ages of 9 and 12 years in the intercorrelations and in the first factor loadings was greater for boys than for girls, are not in conformity with those which have been found by the other investigators mentioned later in this study.

A study by Garrett, Bryan, and Perl (13) was designed to gain information concerning the mental organization of children aged 9, 12, and 15 years. The battery which they used consisted of six memory and four non-memory tests. They found that the mean intercorrelations of all tests for both boys and girls decreased between the ages of 12 and 15 as did the mean intercorrelations of each of their sub-groups of tests for these age levels. Their data, however, show that although this decrease likewise occurred between the ages of 9 and 12 for girls, it did not occur for boys, there being an increase in the mean intercorrelations of all the tests for the latter. This increase also appeared in the mean intercorrelations of all tests by sub-groups except those in which the memory tests predominated.

The multiple factor analysis employed in this study corroborated these findings, showing a consistent decrease with increasing age in the per cent of the variance attributable to the first factor for girls and for boys between the ages of 12 and 15 years. The per cent of variance attributable to the first factor increased in the case of boys aged 9 to 12 years. When only the memory tests were factored this increase disappeared (13, pp. 18-27). Therefore we may presume that had only the non-memory tests been factored a much greater rise would have occurred in the per cent of

variance attributable to the first factor for boys between the ages of 9 and 12 years.

These data were also analyzed by Richards (27) using the centroid method. His solution showed a consistent shift toward discreteness with increase in age for girls. The same general shift was likewise observed for boys but the picture was somewhat confused at age 15 because of the low total variance, and because a third factor was necessary for adequate interpretation at this point.

In addition Richards analyzed Chrysostom's mental test data for children in Grades IV, V, and VI. Here he found a tendency for the clusters to become more discrete in passing from Grade IV to Grade VI. However, in passing from Grade IV to Grade V the factors became less discrete. It was necessary to use three factors at the fifth grade level instead of the two which had sufficed in both the upper and lower grades.

A similar re-analysis of previous studies employing the centroid method was reported by Garrett (11) under whose direction much of the work just cited had been done. In analyzing the data of Schiller, Schneck, Bryan, and two studies by Anastasi, his most important findings were: (*a*) a decrease in the correlation between the verbal and number factors from .825 for Schiller's nine-year-olds to .225 for Schneck's college men and .242 for Anastasi's college woman; (*b*) a correlation of .000 between the memory and number factors and .085 between the memory and verbal factors of Anastasi's college women. His analysis of Bryan's study of young children failed to show any clear-cut structure. It will be recalled that her original data indicated a very close relationship among all of the tests given to her five year old population.

In close agreement with other studies of the mental pattern of adults was the work of Thurstone (35). He administered 56 tests to 240 subjects whose modal age was 18 years. From these tests he obtained 12 orthogonal factors, nine of which he named. These latter have come to be known as the Primary Mental Abilities.

Thurstone's Theory of Independent Mental Abilities was subjected to a test in an experiment conducted by Morrow (23). He gave eight tests consisting of 23 variables to 80 college men and found that, ". . . the correlations are mainly positive . . . although rather low, thereby indicating slight degrees of interrelationships among the abilities tested" (23, p. 84). When he submitted his data to analysis by the centroid method, four factors emerged. The correlation between Factor I and Factor II was .242 and that between Factor I and Factor III was .500. These findings led him to con-

clude that: "It is apparent . . . that the factors are not orthogonal. Therefore, Factors I, II, and III are apparently interrelated" (23, p. 102).

Using Thurstone's well-known test of Primary Mental Abilities, Clark (7) studied changes in the mental pattern of boys aged 11, 13, and 15 years. She found that, with the exception of memory, there was a consistent drop in the intercorrelations of these abilities with increasing age, and that this trend was not influenced by the general intelligence level of the subjects.

Reichard (26), on the other hand, using a battery consisting of three verbal, two number, one spatial, and two memory tests to study the changes in mental organization occurring in boys and girls aged 9, 12, and 15 years, found a consistent increase in her inter-test correlations for both sexes between the 9 and 12 year age levels. The increase for boys was much more marked than was the increase for girls. After the age of 12 years the expected decrease occurred in both sexes. For girls this decrease was so great that the mean intercorrelation for 15-year-olds was below that for 9-year-olds. But in the case of boys the mean intercorrelation for 15-year-olds remained somewhat greater than the mean intercorrelation for 9-year-olds. The average loadings of the first centroid factors which were extracted from the correlation matrices confirmed the trends shown by the mean inter-test correlations.

Recalling the results of the study of Garrett, Bryan, and Perl on these same three age groups, it is apparent that Reichard's results in part confirm and in part contradict those of the former. However, it should also be noted that Reichard's battery was not as heavily loaded with memory tests as was the battery of Garrett, Bryan, and Perl.

An experiment employing the method of factor analysis used by the present investigator, i.e., the bi-factor solution, was conducted by Swineford and Holzinger (34). For their study a battery consisting of three each of verbal, spatial, speed, and memory tests, plus one general test, was administered to 457 seventh and eighth grade boys and girls. One year later 385 of the original group were retested using the same battery plus the Woody-McCall Mixed Fundamentals test. Their results showed that at the seventh and eighth grade level the common factors accounted for approximately 45 per cent of the total variance, and that one year later these same tests produced common factors which accounted for nearly 53 per cent of the total variance. It should be noted that these results exhibit a trend similar to that reported by Reichard, even though a different method of factor analysis was used.

Swineford (33) administered a battery of six tests to 212 pupils in the

eighth grades of seven different schools in the Autumn of 1944, and administered the tests again a year later to these same pupils when they were in the ninth grades. She also tested 173 pupils in the seventh grade in 1944, and retested the same pupils when they were in the ninth grade in 1946. The bi-factor analysis yielded a general and a verbal factor, both of which Swineford concluded could be given the same interpretation at all three grade levels (33, p. 263). She also studied the differentiation of the factor patterns at the three age levels to determine whether or not her data would support the hypothesis that, with increase in age, "individuals tend toward the increasing use of their 'special abilities' with a corresponding decrease in the use of 'general ability'" (33, p. 263). She found no evidence to support the hypothesis, but, on the contrary, found that with increase in age, the general factor increased in both its absolute and relative contribution to the total test variance (33, p. 271).

The relationship between general and group factors has also attracted the attention of workers in other fields of ability. For instance, Babcock and Emerson (2) found a consistent increase in the mean rank order correlations between six of the seven sub-tests of the MacQuarrie test of Mechanical Ability and the Terman Vocabulary Test for subjects between the ages of 14 and 17 years, and also between the ages of 17 years and adulthood (20 to 28 years). In the one remaining sub-test there was a decrease in the mean rank order correlations between the ages of 14 and 17, but none between the latter mentioned age groups.

The experiment that covers the largest age range was that of Balinsky (13), who studied six different groups ranging in age from 9 to 59 years. His subjects were selected from those who had been used for the standardization of the *Wechsler-Bellevue Intelligence Scale*. Several interesting findings were reported, among which were: (*a*) a general factor common to all the tests appeared at the age of 9, became submerged during the intervening years, then reappeared in the 50 to 59 year group; (*b*) the average of the sub-test intercorrelations decreased from age 9 to the 25 to 29 year age group, and then progressively increased up to the 50 to 59 year age group; (*c*) the percentage of variance attributable to the first factor decreased from 38 per cent at age 9 to 20 per cent at the 25 to 29 age level, and then increased again to 45 per cent at the 50 to 59 age level; (*d*) there was a definite tendency for sub-tests to be found in different clusters at different age levels.

A critical examination of the studies reported in this section revealed a general agreement that there was a differentiation in the factor pattern of girls after the age of nine years, even though this differentiation was not

as marked in those studies in which memory tests played a minor rôle as it was in the studies using batteries more heavily loaded with such tests. In the case of boys the differentiation of the factor pattern usually did not become apparent before the age of 12 years unless the batteries were quite heavily loaded with memory tests. In those studies in which verbal, number, and spatial tests played a more important rôle, the results were both confusing and inconclusive. Therefore an independent test of the differentiation hypothesis using non-memory tests seemed to be justified.

C. Previous Work on the Effect of Test Difficulty

The usual method of studying shifts in the factor pattern is to use the same tests for subjects of all ages (29, p. 59). It is evident that this technique does not control the difficulty of the test material as judged by the mean scores of the subjects, but it is tacitly assumed that this difference in mean scores is merely an additive factor which has no effect upon the intercorrelations of the tests, or upon their factorial patterns, provided the tests have sufficient range to insure that the scores for the different age groups tested are not significantly skewed.

However, a review of the research concerning the effects of test difficulty both upon inter-test correlations and factorial patterns revealed several interesting and significant findings.

Hertzman (17) analyzed data obtained by Smith (32) on 180 college men using four numerical and three spatial tests. He found that by dividing each of the tests into two parts, an easy and a difficult part, and by scoring the subjects separately upon each of these parts, the resulting inter-test correlations were different. He concluded that: "Other factors being equal, differences in the level of the same kind of material result in a depression of the correlation between samples of it . . . the easy and the difficult items consequently cannot be thought of as measuring exactly the same ability" (17, p. (64).

Observation of the fact that "As difference between two levels of difficulty increases, there is a drop in the correlation which measures relationship between performance at the two levels" (28, p. 5), led Sargent to conduct a rather ingenious piece of research. He trained 15 college students, all of whom were majoring in psychology, to "talk out" the thought processes they engaged in while solving problems of varying degrees of difficulty. In addition to his work with these students he tested an additional 50 college students individually, and 317 college students using the group method. In all cases the test employed was one composed of 40 disarranged words classified as easy, medium, and difficult. His conclusion was that, "Hereto-

fore it has been assumed that easy and difficult levels of a problem differ only in degree; these results suggest that, in terms of mental processes involved, they tend to be different kinds of problems" (28, p. 54).

Studies of the effect of difficulty upon factorial pattern were made by Ferguson (10) and by Guilford (15). Ferguson showed that, theoretically, tests which were similar in difficulty should correlate more highly than tests separated by wide differences in difficulty level. He made an empirical test of this theory by administering six difficulty levels of a *Moray House Verbal Intelligence Test* to 11-year-old children. The first factor loadings of each test tended to decrease as the tests became either more or less difficult than the two middle tests of the battery. The second factor had positive loadings on the easy sub-tests, approximately zero loadings on tests of mean difficulty, and negative loadings on the sub-tests of greater than mean difficulty. In direct support of his theory he found that at each point in the scale, tests of approximately equal difficulty were more nearly alike in their factorial composition than were the tests separated by wide differences of difficulty.

Guilford (15) analyzed the 10 sub-tests of the Seashore test of pitch discrimination. He found that the loadings of the one factor which accounted for the greatest share of the variance decreased systematically with increasing difficulty. Two other factors were found, but we are not immediately concerned with their interpretation. From this study we may infer that the first factor loading on the easier forms of tests should be higher than the first factor loadings on the more difficult forms of the same material, other things being equal.

Three important conclusions can be drawn from the results of these studies. First, differences in the difficulty levels of tests directly affect both the intertest correlations and the factor pattern which emerge from them. Second, minimizing the differences in difficulty among tests increased their intercorrelations, and hence the observed degree of relationship among them. Third, the same item measures one ability on one level of difficulty, and a different ability on a different level of difficulty.

In view of these facts, the previously mentioned "tacit assumption" of many investigators that using the same form of one test for more than one age level will not directly affect inter-test correlations and factor patterns is here questioned, because scores obtained on a test by younger subjects are not obtained from the same difficulty levels as are scores obtained by older subjects on the same test. When children of two or more different ages are given the same test, the scores of the older subjects consist of correct responses to:

(*a*) Items which for them were much easier than were any of the items for the younger subjects; (*b*) items which for them were of average difficulty as was an entirely different group of items for younger subjects; (*c*) a small number of items which for them were relatively very difficult as was a much larger number of items for the younger subjects.

A consideration of all of the above facts led the investigator to consider whether or not minimizing the differences in the relative difficulty of tests for two populations of different ages might not increase the observed degree of relationship between the two age groups.

D. Design of the Experiment

The review of the literature dealing with changes in mental organization with age seemed to justify an independent test of Professor Garrett's hypothesis that "Abstract or symbol intelligence changes in its organization as age increases from a fairly unified and general ability to a loosely organized group of abilities or factors" (12, p. 373). In addition, a study of the literature dealing with the effect of differences of difficulty level upon the intercorrelations of tests established the need for testing the hypothesis of the investigator that: The observed differences between the factor patterns for two age groups are due not only to the age differences of the groups in question but also to the differences in the difficulty levels of the test materials for the two populations.

The experiment reported in this study was specifically designed for the purpose of testing these two hypotheses. Each of the tests to be used was to consist of two forms which were similar in content but which differed in difficulty level. Form I was to be approximately as difficult for the younger children as was Form II for the older children. The two groups selected were to be well separated in age, of the same sex, and homogeneous with respect to racial and cultural background. Each form of each test was to be administered to both the younger and the older age groups, and the results were to be submitted to a factorial analysis using the bi-factor method.

By the first hypothesis, the factor pattern of the older subjects should, form by form, exhibit greater differentiation than should the factor pattern of the younger subjects for the same forms.

By the second hypothesis, the differentiation of the factor pattern of the younger age group on Form I should be more nearly like that of the older subjects on Form II than like that of the older subjects on Form I, and the degree of differentiation of the younger age group on Form II should be least like that of the older age group on Form I.

II. THE EXPERIMENT

A. The Population

The younger subjects used in this experiment were 103 white boys aged 9 years to 9 years 11 months. The older subjects were 115 white boys aged 12 years to 12 years 11 months. They were all regularly enrolled in four elementary schools in Springfield, Illinois.

The subjects may be described as a cross section of the school population of a typical mid-western city, excluding the exceptionally under-privileged and foreign speaking groups. All were able to speak, understand, read, and write simple English. A careful check of the public school records showed that there was no adverse selection due to the more brilliant boys having been promoted to another school. By testing 9 and 12-year-old boys, and by conducting the program in a system in which the children were kept in the same building from the kindergarten through the first eight grades, all persons within the age ranges regularly enrolled in the public schools were reached. The few that were absent or turned in defective papers were excluded from the experimental group.

It is possible that this experimental population was slightly above the national average in intelligence. Recently obtained scores which were available for the fourth grade in the system showed that pupils in the four schools tested tended to be somewhat superior in ability. A group that was slightly superior in ability but homogeneous with respect to race and cultural background was to be preferred in this study to a more nearly normal but heterogeneous sample (19, pp. 24-33). The standard deviations of the scores obtained indicate that the range of ability was not seriously truncated.

B. Construction of the Tests

Verbal, number, and spatial materials were selected for test content because the character of the changes in the relationships among these factors between the ages of 9 and 12 years seemed to be more uncertain than were the changes between any one of them and memory.

Each test was constructed by first assembling a common pool of items for it. In some cases bi-serial coefficients of correlations between the several items and the common pool were available. In other cases the pool was made up of a collection of items from a test which had been subjected to careful study by its author. For approximately half of the tests, there were data showing the per cent of typical 9-year-old children and of typical 12-

year-old children who had on other occassions passed each item. In some cases the items had been scaled by the authors of the tests from which they were drawn and these data were available. In all cases enough was known about the difficulty of the items to enable the investigator to place them in the 9-year-old form or the 12-year-old form on the basis of their difficulty. The items comprising the 10 tests used in this study were gathered from four sources.

(*a*) Professor F. H. Finch of the College of Education of the University of Illinois very generously made available five groups of unpublished, experimental items. For many of these items the bi-serial coefficient of correlation between the several items and their respective pools were available. For each one, there was accurate information concerning its difficulty for younger and older children. The difference in the difficulty level for children separated in age furnished an indication of each item's ability to discriminate between younger and older subjects.

From these five groups, carefully matched items were drawn to make two forms each of the following tests: "Finding Middle Number," "Number Series," "Number Patterns," "Figure Exercises," and "Boys and Girls."

(*b*) The University of Chicago Press granted permission to use items drawn from the "Chicago Mental Growth Battery," by Freeman and Wenger. All items in this battery are scaled on the basis of difficulty. From this source items were drawn to make two forms of the "Paper Form Board," the "Vocabulary," the "Opposites," and the "Word Grouping" tests. The preliminary trial, to be described later, showed that the "Paper Form Board" was not satisfactory for this experiment, so it was dropped.

(*c*) The World Book Company granted special permission to reproduce and use the items in the Figure Dividing Section of the "Pintner General Ability Tests: Non-Language Series, Forms *K* and *L*." These items were supplemented by a group of 24 similar items constructed by the experimenter, and from this combined pool, items were drawn to make two forms of the "Figure Dividing" test.

(*d*) The remaining test entitled "Arithmetic Reasoning" was developed by the experimenter after a careful analysis of the types of verbal problems employed in standard arithmetic texts.

It will be recalled that the design of this experiment necessitated the use of one form of each test which would be as difficult for younger subjects as was the second form for older subjects. The two forms of each test were constructed by alternately drawing items for those tests from their respective pools. An item of a given difficulty level for boys aged nine years

was drawn for Form I, then an item of the same difficulty level for boys aged 12 years was drawn for Form II. This method of drawing items more nearly assured homogeneity of content and of difficulty than might have been secured had either of the forms been made first.

The system employed in distributing item difficulty in each form was based upon a rough approximation to the normal curve. If the two forms were to contain 30 items each, the first item in Form I would have a "p," per cent passing, of 90 or above for children aged nine, the first in Form II would have a "p" of 90 or above for children aged 12; two items in each form would have "p's" between 80 and 90; five or six items in each form would have "p's" between 60 and 80; 12 to 14 items in each form between 40 and 60; five or six items in each form between 20 and 40; and two or three items in each form would have "p's" below 20.

In all probability this method of distributing item difficulty had a great deal to do with the general tendency of the distributions of total scores to be platykurtic. It is apparent that in such a distribution of difficulty, approximately two-thirds of the items were within 20 per cent of the 50 per cent difficulty point. Cook's (8) investigation of the effect of the distribution of word difficulty upon the shape of the distribution of total spelling scores showed that concentrating words at the 50 per cent tended to produce a flat distribution.

C. Description of the Tests

1. *Finding Middle Number*

The items in this test consisted of five or seven numbers in a disarranged order. The subjects were instructed to "find the middle-sized number" or the one which "if the numbers were in order from the smallest to the largest . . . would be in the middle," and to fill in the space under it. In Form I, the five numbers in each item varied in size from one to four digits. In Form II, 15 items consisted of five numbers varying in size from three to five digits, nine items consisted of seven numbers of three or four digits each, and one item was defective because it inadvertently contained six, an even number, of parts. Number of items in each form (including the defective item in Form II), 25; working time for each form, 7 minutes.

2. *Number Series*

This test was in multiple choice form, the correct answer being one of the five numbers in the adjacent column to the right, headed "next number." The subjects were directed to "find the rule" for each row, then to use this

rule "to figure out what number should come next, and to fill in the space under it." Since they were required to use one of the five choices in the second column, ambiguities could be eliminated by failing to include among the answers any numbers that could be obtained by alternate solutions. Number of items in each form, 26; working time for each form, 10 minutes.

3. *Number Patterns*

This test was essentially an adaptation of the number series problem. In the problem of Form I, a row of two numbers, and a second row with a missing number were placed in a square. The subjects were shown how to solve the problem by starting with the first number in a row and figuring out what mathematical computation was performed to get the second number in the same horizontal row. They were then shown that doing "the same thing" to the first number in the second row would give them the missing number for the answer, which was always one of the five numbers below the square. The problem was varied in Form I by using both addition and subtraction and by changing the position of the required answer. In Form II, all four of the fundamental arithmetic processes were used, the position of the required answer was changed, the pattern was increased to nine numbers, and varying numbers of figures in the horizontal rows were omitted. Number of items in each form, 30; working time for each form, 10 minutes.

4. *Figure Exercises*

This test was essentially spatial analogies. Three figures and a blank space were enclosed in a square. There were five figures below the square, one of which belonged in the blank space. By means of the practice exercise and the verbal explanation, the subjects were led to find the relationship between the first and the second figure, and to use this relationship, starting with the third, to find the fourth figure. The correct figure for the blank space was always the one which bore the same relation to the third as did the second to the first. Number of items in each form, 24; working time for each form 8 minutes.

5. *Figure Dividing*

This test consisted of horizontal rows of drawings. In each row, there was first a closed figure, then four line drawings, then a group of pieces which had been cut from the first figure. One of the line drawings would

cut the first figure into the pieces shown in the last group. The subjects were directed to fill in the blank space under the line drawing which would cut the first figure so as to make it look like the last one. Number of items in each form, 30; working time for each form, 8 minutes.

6. *Boys and Girls*

The "Boys and Girls" were drawings of small, doll-like figures of children, the first figure in each row being separated from the rest by a dotted line. The subjects were directed to "look at the first picture to see how the arms are held," then to "make a mark" on that one of the five remaining pictures in the row which held the arms in the same position as did the first child. The figures were both revolved in the plane of the paper and rotated on their own major axes. Thus the first figure might be standing upright, facing the reader and holding the right arm above the head. The correct answer might be a figure which was turned so that the head was down and the back was to the reader, but it would be the only one in the row holding the right arm above the head. Number of items in each form, 27; working time for each form, 9 minutes.

7. *Vocabulary*

This test was made up of rows of six words each, a single word being placed to the left of a vertical line and the five other words or phrases to the right of the line. The subjects were directed to fill in the space under the word or phrase to the right of the line which had the same meaning as the word to the left of the line in that row. Number of items in each form, 27; working time for each form, 9 minutes.

8. *Vocabulary*

This test was made up of rows of six words each, a single word being placed to the left of a vertical line and the five other words or phrases to the right of the line. The subjects were directed to fill in the space under the word or phrase to the right of the line which had the same meaning as the word to the left of the line in that row. Number of items in each form, 30; working time for each form, 6 minutes.

9. *Opposites*

The form of the "Opposites" test was much like that of the Vocabulary test, the only difference being that the subjects were directed to mark the word whose meaning was the exact opposite of the key word located to the

left of the vertical line. Number of items in each form, 30; working time for each form, 6 minutes.

10. *Word Grouping*

In this test each item was made up of five words or phrases. The subjects were told that four of these words or phrases belonged together in some way, but that one did not belong with the rest. They were directed to fill in the space beneath the latter word or phrase. Number of items in each form, 30; working time for each form, 8 minutes.

11. *Arithmetic Reasoning*

In constructing this test, several series of elementary arithmetic texts were examined to determine the types of verbal problems boys aged nine and boys aged 12 years were customarily required to work. The same types of problems, using very simple number combinations and relatively simple language, were used. The subjects were permitted to use an extra sheet of paper on which to work their problems if they wished. Number of items in each form, 30; working time for each form, 10 minutes.

12. *Paper Form Board*

This test was unsatisfactory and was replaced by the "Boys and Girls" test after the preliminary trial.

The tests were assembled into 10 booklets. The cover page provided space for the subject's name, sex, grade, school, city, age, and date of birth. The cover page also contained examples and whatever directions the subject was to read prior to starting the test. The booklet for each test was so assembled that Form I followed the cover page, face up, and Form II followed Form I, back to back.

The directions for working both forms were the same for all tests throughout the entire battery. The subjects were told to continue working through the pages of Form I until they reached the blank page in the middle of the booklet. This blank page was, of course the back of the last page of Form II of the test. They were directed to check their work on Form I and after checking, to wait quietly for further directions. Upon the expiration of the allotted time the subjects were asked to stop, close their test booklets so that their names would be on top, turn the whole booklet over, and start working on the second part, continuing until they again reached the blank page in the middle. They were directed to check their answers to the questions in this part of the test, but were told not to turn back to the material in the first part of the test.

D. Preliminary Trial

To verify difficulty levels of the two forms as well as their ability to discriminate between older and younger subjects, a preliminary trial of the test materials was held in the elementary school of Staunton, Illinois, during the last week of February, 1947. The test included all children in grades four through eight, and in addition, children who had reached their ninth birthdays, but who were still in grades lower than the fourth. Complete notes were taken on such matters as optimum working times, desirable changes in the manuals, typographical errors in the mimeographed materials, and subjects' attitudes toward the different tests.

The tests were then carefully scored. On the basis of the mean scores of boys and of girls at each age level, and the notes taken at the time of testing, manuals and tests were revised. The "Paper Form Board" test was eliminated from the battery, because it did not discriminate between the two age groups. The test, "Boys and Girls," was developed then to replace the "Paper Form Board." Several of the other tests were revised to make Form I more nearly of 50 per cent mean difficulty for boys nine years of age and Form II of more nearly 50 per cent mean difficulty for boys aged 12. Minor revisions were made in the manuals and the time limits were established.

E. Testing Procedure

The revised tests were administered in the Springfield schools between April 22 and April 29, 1947. Three sessions of approximately 90 minutes each were required to complete the battery. In every school, the work was scheduled so that each age group had one session in the morning and another the same afternoon. The third session was held within 48 hours of the other two. At each session at least one test involved numbers, another verbal material, and a third spatial material. The various tests were given in a different order to each group, and the order was so varied that it could have had little systematic effect upon the scores of the subjects.

The general testing conditions were excellent. Besides the examiner in charge, an adult assistant was present at all times in the larger groups. All examinations were conducted by the supervisor of testing in the Springfield schools or by the experimenter. The usual precautions were taken to prevent copying. The subjects were told that they were part of an experiment and that although the scores on the tests would not affect their school marks, their scores would be returned to them so they could see how well they had done. The general atmosphere was conducive to good work at all times.

F. Reliability and Inter-Form Correlation

The reliabilities of the tests as measures of the abilities of the age groups studied in this experiment were not known prior to their use with the experimental population. This information is particularly important in factor studies, because the factorial analysis of unreliable measures makes the meaning of the factors extracted extremely uncertain.

Guttman's (16) lower bound, L_3, was used as the measure of reliability. The formula involves a summation of the item variances as well as the variance of the total test. To obtain the item variances the papers were so scored that the per cent passing each item could be computed. The lower bound of reliability of Form I, Form II, and of the Combined Form, was computed for the experimental population of 9-year-old boys and of 12-year-old boys, separately. These lower bounds are reported in Table 1 for the younger group, and in Table 2 for the older group.

TABLE 1

Lower Bounds of Reliability for Boys Aged 9 Years

Test	Form I	Form II	Combined Form
1	.934	.756	.929
2	.926	.875	.944
3	.961	.761	.936
4	.910	.901	.947
5	.939	.888	.950
6	.805	.883	.912
7	.919	.700	.914
8	.848	.638	.866
9	.884	.749	.909
10	.867	.741	.898

Read: For test 1, the lower bound of reliability of Form I was .934; of Form II, .756; of the Combined Form, .929; etc.

Names of Tests

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

It may be noted that the reliabilities of each form based upon data from each age group are generally as high as those ordinarily reported for psychological tests. The lowest bound for subjects aged nine years is that of Form II of Test 8, which is .638. The lowest bound for any of the tests at the 12-year level is .593, that of the reliability of Form II of Test 7.

The correlations between the two forms of each of the 10 tests used in this

TABLE 2

LOWER BOUNDS OF RELIABILITY FOR BOYS AGED 12 YEARS

Test	Form I	Form II	Combined Form
1	.955	.868	.956
2	.950	.946	.971
3	.967	.915	.961
4	.919	.909	.951
5	.896	.838	.918
6	.884	.942	.954
7	.866	.593	.866
8	.826	.674	.858
9	.823	.758	.876
10	.872	.819	.918

READ: For test 1, the lower bound of reliability of Form I was .955; of Form II, .868; of the Combined Form, .956; etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

TABLE 3

CORRELATION OF FORM I AND FORM II OF EACH OF THE TEN TESTS

Test number	rI II Boys aged 9 years	rI II Boys aged 12 years
1	.657	.769
2	.749	.845
3	.230	.531
4	.811	.827
5	.736	.709
6	.749	.797
7	.704	.714
8	.697	.673
9	.784	.744
10	.747	.841

READ: For test 1, the correlations between Form I and Form II for boys aged nine years was .657, for boys aged 12 years .769, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

study are given in Table 3. For boys aged nine years, seven of the correlations were larger than .70, two were between .65 and .70, and one, that for Test 3, was .230. The second form of this test was extremely dif-

ficult for this group and this fact no doubt contributed to the low correlation. For boys aged 12 years, the correlations between the two forms exceeded .70 in eight of the 10 tests. Of the two remaining tests one had a correlation of .673 between the two forms. The correlation in the case of Test 3 had risen some 30 points, now being .531.

It should be pointed out again that these are not equivalent forms in the sense in which this term is generally used. They are two forms having similar content that were intended to differ in difficulty. We should therefore expect these correlations to be somewhat lower than would have been the case had the forms been equivalent.

III. DESCRIPTION OF THE DATA

The following discussion will contain a very complete and detailed description of the data obtained in this study. This is essential because of the general inadequacy of existing error formulae for factor coefficients and residuals. Holzinger's bi-factor solution is the only method of factor analysis for which such error formulae exist and he very strongly advises that they be used "with great caution and with a full realization of the underlying assumptions," many of which are only roughly satisfied in most empirical data (18, p. 131). Even if adequate error formulae existed, we still would know all too little of the conditions under which factors emerge because factors certainly are not entities independent of the conditions in which they are found. Kelley has expressed this point well when he said: "There is no search for timeless, spaceless, populationless truth in factor analysis; rather, it represents a simple, straightforward problem of description in several dimensions of a definite group functioning in definite manners, and he who assumes to read more remote verities into the factorial outcome is certainly doomed to disappointment" (20, p. 120). Essentially the same idea was expressed by Garrett who stated, ". . . factors (are) . . . (*a*) experimentally determined, (*b*) operationally defined, and (*c*) subject to tests of validity" (11, p. 295).

A. The Distribution of Scores

The distributions of the scores for boys aged nine years and for boys aged 12 years on each form of each of the 10 tests are described in terms of their respective means, standard deviations, and third and fourth moments expressed in standard units. The first two descriptive constants are commonly used in educational and psychological literature. Since the third and fourth moments are not as familiar, a brief explanation of each may be helpful.

When the third moment is taken about the mean, it is a measure of skewness (21, p. 71). If the distribution is symmetrical, its value is zero. If the distribution is not symmetrical its sign indicates the direction of skewness. The third moment expressed in standard units is denoted by the Greek letter alpha and its symbol is "a_3". It is defined as the mean of the third powers of the deviations from the mean of the distribution divided by the third power of the standard deviation. Since a_3 is independent of the type of distribution and of the unit of measurement, it is a more satisfactory measure for comparing the symmetry of the distributions in this experiment than are the more commonly used measures.

The fourth moment taken about the mean is a measure of kurtosis (21, p. 71). When expressed in standard units it, too, is denoted by alpha, its symbol being "a_4". By definition, a_4 is the mean of the fourth powers of the deviations taken from the mean, divided by the fourth power of the standard deviation of the distribution. Since, like a_3, it is independent of both the shape of the distribution and of the unit of measurement, it too is more suitable for use in this study than are the more commonly employed measures. Its value for the normal curve is three. If a_4 is greater than three, the distribution is leptokurtic; if less than three, it is platykurtic.

To facilitate comparisons, the constants of the distribution for the two age groups on the different forms of the tests have been assembled in separate tables. It is not possible to compare the means and standard deviations of the various tests directly because these are in terms of raw scores. However, the measures of skewness and kurtosis are directly comparable because they are expressed in terms of the appropriate powers of the standard deviations of their respective distributions.

These data for nine year old boys on Form I, Form II, and the Combined Form are presented in Tables 4, 5, and 6, respectively. Table 4

TABLE 4

MEAN, STANDARD DEVIATION, a_3 AND a_4 OF THE DISTRIBUTION OF SCORES FOR BOYS AGED 9 YEARS ON FORM I OF EACH OF TEN TESTS

Test	Mean	Standard Deviation	Skewness a_3	Kurtosis a_4
1	9.68	7.04	0.50	2.08
2	11.51	7.11	0.39	1.88
3	13.03	9.96	0.40	1.65
4	9.88	6.50	0.79	2.36
5	22.17	7.40	—1.15	3.09
6	11.31	4.80	0.59	2.86
7	17.74	7.03	—0.34	1.80
8	15.87	5.16	—0.25	2.41
9	17.03	6.56	—0.26	2.07
10	9.76	4.95	0.25	2.26

READ HORIZONTALLY: The mean of Test 1 is 9.68, its standard deviation is 7.04, skewness (a_3) is positive 0.50, and its kurtosis (a_4) is 2.08, which indicates that it is flatter than the normal curve.

READ VERTICALLY: The mean of Test 1 is 9.68, of Test 2, 11.51, etc.; the standard deviation of Test 1 is 7.04, of Test 2, 7.11, etc.; the skewness (a_3) of Test 1 is 0.50, of Test 2, 0.39, etc.; and the kurtosis (a_4) of Test 1 is 2.08, of Test 2, 1.88, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

TABLE 5

MEAN, STANDARD DEVIATION, a_3 AND a_4 OF THE DISTRIBUTION OF SCORES FOR BOYS AGED 9 YEARS ON FORM II OF EACH OF TEN TESTS

Test	Mean	Standard Deviation	Skewness a_3	Kurtosis a_4
1	4.34	3.34	0.86	3.06
2	5.54	4.95	1.14	3.78
3	3.82	3.35	1.43	5.04
4	9.49	6.17	0.70	2.35
5	16.43	6.67	—0.29	2.43
6	6.94	5.37	1.54	5.07
7	9.81	3.72	—0.03	3.02
8	9.31	3.55	0.11	2.67
9	9.61	4.52	0.48	2.96
10	3.77	2.83	0.64	2.82

READ HORIZONTALLY: The mean of Test 1 is 4.34, its standard deviation, 3.34, skewness (a_3) is positive 0.86 and its kurtosis (a_4) is 3.06, which indicates that it is approximately normal.

READ VERTICALLY: The mean of Test 1 is 4.34, of Test 2, 5.54, etc.; the standard deviation of Test 1 is 3.34, of Test 2, 4.95, etc.; the skewness (a_3) of Test 1 is 0.86, of Test 2, 1.14, etc.; and the kurtosis (a_4) of Test 1 is 3.06, of Test 2, 3.78, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

shows that four of the distributions were negatively skewed and that six were positively skewed. Nine of the curves were platykurtic and the remaining one was practically normal. Examining the constants for Form II recorded in Table 5 we find that there has been a decided shift toward positive skewness, there now being eight distributions so skewed and only a relatively slight negative skewness in the other two. Also we find that in comparison with Form I, eight of the distributions were more leptokurtic, that of test five was more platykurtic, and that of test four showed practically no change.

The constants for the scores on the Combined Form shown in Table 6 reveal the expected general increase in the standard deviations. Six of the distributions were positively skewed, three showed slight negative skewness, and that of Test 5 showed marked negative skewness. Eight of the curves were definitely platykurtic, Test 5's was practically normal, and the curve of Test 6 was leptokurtic.

Similar presentations of the data for 12-year-old boys on Form I, Form II, and the Combined Form are to be found in Tables 7, 8, and 9, respectively. The distributions of scores for all tests in Form I reported in Table 7 showed negative skewness. Half of the distributions were leptokurtic and the other half

TABLE 6

MEAN, STANDARD DEVIATION, a_3 AND a_4 OF THE DISTRIBUTION OF SCORES FOR BOYS AGED 9 YEARS ON COMBINED FORM OF EACH OF 10 TESTS

Test	Mean	Standard Deviation	Skewness a_3	Kurtosis a_4
1	14.02	9.57	0.66	2.49
2	17.06	11.28	0.74	2.59
3	16.84	11.22	0.55	2.30
4	19.38	12.05	0.81	2.43
5	38.60	13.11	—0.88	2.94
6	18.25	9.51	1.12	4.06
7	27.54	10.00	—0.22	1.88
8	25.18	8.05	—0.10	2.45
9	26.56	10.61	—0.12	2.22
10	13.52	7.31	0.37	2.51

READ HORIZONTALLY: The mean of Test 1 is 14.02, its standard deviation is 9.57, skewness (a_3) is positive 0.66, and its kurtosis (a_4) is 2.49 which indicates that it is flatter than the normal curve.

READ VERTICALLY: The mean of Test 1 is 14.02, of Test 2, 17.06, etc.; the standard deviation of Test 1 is 9.57, of Test 2, 11.28, etc.; the skewness (a_3) of Test 1 is 0.66, of Test 2, 0.74, etc.; and the kurtosis (a_4) of Test 1 is 2.49, of Test 2, 2.59, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

TABLE 7

MEAN, STANDARD DEVIATION, a_3 AND a_4, OF THE DISTRIBUTION OF SCORES FOR BOYS AGED 12 YEARS ON FORM I OF EACH OF 10 TESTS

Test	Mean	Standard Deviation	Skewness a_3	Kurtosis a_4
1	16.68	7.82	—0.71	2.07
2	18.06	7.67	—0.78	2.24
3	22.11	9.25	—1.00	2.48
4	14.37	6.75	—0.23	1.73
5	25.97	4.63	—2.30	7.91
6	17.52	6.02	—0.37	2.14
7	24.77	4.49	—1.68	5.63
8	21.95	4.62	—1.42	5.84
9	23.24	4.71	—1.33	5.14
10	19.86	5.40	—0.65	3.43

READ HORIZONTALLY: The mean of Test 1 is 16.68, its standard deviation 7.82, skewness (a_3) is negative 0.71 and its kurtosis (a_4) is 2.07, which indicates that it is flatter than the normal curve.

READ VERTICALLY: The mean of Test 1 is 16.68, of Test 2, 18.06, etc.; the standard deviation of Test 1 is 7.82, of Test 2, 7.67, etc.; the skewness (a_3) of Test 1 is negative 0.71, of Test 2, negative 0.78, etc.; the kurtosis (a_4) of Test 1 is 2.07, of Test 2, 2.24, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

TABLE 8

MEAN, STANDARD DEVIATION, a_3 AND a_4, OF THE DISTRIBUTION OF SCORES FOR BOYS AGED 12 YEARS ON FORM II OF EACH OF 10 TESTS

Test	Mean	Standard Deviation	Skewness a_3	Kurtosis a_4
1	9.25	5.12	0.04	2.26
2	12.72	7.92	—0.07	1.66
3	9.82	6.79	0.80	3.49
4	13.56	6.41	—0.38	1.91
5	21.50	5.20	—1.27	4.55
6	13.27	8.03	0.19	1.70
7	13.58	3.10	—0.72	3.63
8	14.03	3.69	—0.30	2.93
9	15.00	4.85	0.07	2.76
10	11.10	4.37	—0.47	2.60

READ HORIZONTALLY: The mean of Test 1 is 9.25, its standard deviation is 5.12, skewness (a_3) is positive 0.04 and its kurtosis (a_4) is 2.26, which indicates that it is flatter than normal.

READ VERTICALLY: The mean of Test 1 is 9.25, of Test 2, 12.72, etc.; the standard deviation of Test 1 is 5.12, of Test 2, 7.92, etc.; the skewness (a_3) of Test 1 is positive 0.04, of Test 2, negative 0.07, etc.; the kurtosis (a_4) of Test 1 is 2.26, of Test 2, 1.66, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

TABLE 9

MEAN, STANDARD DEVIATION, a_3 AND a_4, OF THE DISTRIBUTION OF SCORES FOR BOYS AGED 12 YEARS ON COMBINED FORM OF EACH OF 10 TESTS

Test	Mean	Standard Deviation	Skewness a_3	Kurtosis a_4
1	25.93	12.20	—0.45	1.96
2	30.78	14.97	—0.41	1.82
3	31.93	14.08	—0.33	2.20
4	27.83	12.51	—0.28	1.79
5	47.48	9.05	—1.87	6.58
6	30.79	13.33	0.04	1.83
7	38.36	7.05	—1.43	5.11
8	35.98	7.61	—1.10	4.88
9	38.16	8.84	—0.70	3.41
10	30.97	9.38	—0.59	2.98

READ HORIZONTALLY: The mean of Test 1 is 25.93, its standard deviation 12.20, skewness (a_3) is negative 0.45 and its kurtosis (a_4) is 1.96, which indicates that it is flatter than the normal curve.

READ VERTICALLY: The mean of Test 1 is 25.93, of Test 2, 30.78, etc.; the standard deviation of Test 1 is 12.20, of Test 2, 14.97, etc.; the skewness (a_3) of Test 1 is negative 0.45, of Test 2, negative 0.41, etc.; and the kurtosis (a_4) of Test 1 is 1.96, of Test 2, 1.82, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

were somewhat platykurtic. On Form II, as shown in Table 8, we find that the distributions were much more symmetrical, and that seven of them were platykurtic while the remaining three were slightly leptokurtic. On the Combined Form the expected increase in the standard deviations again concurred. Table 9 shows that there was a general tendency toward slight negative skewness and that six of the distributions tended to be platykurtic, whereas four tended to be leptokurtic.

B. The Adequacy of the Tests

The data in Table 10 may be examined to determine how well these tests fulfilled the requirements set up in Section II. The ratios of the differences between the specified means to their respective standard errors are indicated in the several columns of this table. For purposes of this discussion a difference will be considered statistically significant when its entry exceeds 2.34. This ratio establishes the significance of the difference at approximately the one per cent confidence level. This is true because we are concerned in this study with positive differences, only.

TABLE 10

Ratio of Differences between Selected Means to Their Respective Standard Errors

Test	Form I, Age 9 Minus Form II, Age 9	Form I, Age 12 Minus Form II, Age 12	Form I, Age 12 Minus Form I, Age 9	Form II, Age 12 Minus Form II, Age 9
1	9.93	15.70	6.96	8.47
2	12.81	13.15	6.54	8.11
3	9.59	16.37	6.96	8.41
4	1.04	2.23	4.99	4.78
5	11.31	12.66	4.49	6.22
6	12.17	9.37	8.46	6.90
7	15.65	38.20	8.70	8.09
8	17.98	24.50	9.12	9.62
9	18.55	25.84	7.96	8.49
10	6.66	32.08	14.41	14.88

Read Horizontally: For Test 1, the difference between the means of the scores on Form I for boys aged 9 years and on Form II for boys aged 9 years was 9.93 times its standard error; the difference between the means of the scores on Form I for boys aged 12 years and on Form II for boys aged 12 years was 15.70 times its standard error; etc.

Read Vertically: To secure for each of the tests the ratio between any desired difference and its respective standard error, read down the appropriate column.

Names of Tests

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

TABLE 10 B

RATIO OF DIFFERENCES BETWEEN SELECTED MEANS TO THEIR RESPECTIVE STANDARD ERRORS

Test	Comb., Age 12 Minus Comb., Age 9	Form II, Age 12 Minus Form I, Age 9	Form I, Age 12 Minus Form II, Age 9
1	8.06	—0.51	15.42
2	7.69	1.18	14.45
3	8.79	—2.75	19.82
4	5.08	4.19	5.58
5	5.75	—0.77	12.15
6	8.05	2.21	13.71
7	9.13	—5.53	26.92
8	10.15	—3.00	22.77
9	8.71	—2.57	21.59
10	15.39	2.12	27.94

READ HORIZONTALLY: For Test 1, the difference between the means of the scores on the Combined Form for boys aged 12 years and the Combined Form for boys aged 9 years was 8.06 times its standard error; the difference between the means of the scores on Form II for boys aged 12 years and on Form I for boys aged 9 years was —0.51 times its standard error; etc.

READ VERTICALLY: To secure for each of the tests the ratio between any desired difference and its respective standard error, read down the appropriate column.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

The first requirement was that Form I should be easier than Form II for both groups. An inspection of the second and third columns in Table 10 shows that for boys aged nine years and for boys aged 12 years Form I of each of the 10 tests was easier than Form II. In nine of the tests this difference was established with practical certainty. In Test 4, however, the difference was not statistically significant.

The second requirement was that each test in each form be easier for 12-year-old boys than for 9-year-old boys. The differences between the mean achievements of the two age groups on the tests in Form I and Form II are given in the third and fourth columns, respectively. The corresponding differences on the Combined Form are given in the second column of Table 10B. In every case the mean of the scores for boys aged 12 years was significantly higher than for boys aged nine years. Therefore the tests fulfilled the second requirement.

The third requirement was that Form I of each test be approximately as difficult for boys aged nine years as Form II was for boys aged 12 years. By checking the signs of the differences recorded in the third column of

Table 10B it will be noted that for Form I, six of the 10 tests were somewhat easier for the younger boys than was Form II for the older ones. In five cases the difference in relative difficulty was not statistically significant. Of the five significant differences, four favored the 9-year-olds and one the 12-year-olds. Thus we find that the third requirement was rather well fulfilled.

In view of the fact that reference will later be made to the differences between the factor patterns of boys aged 12 years on Form I and of boys aged nine years on Form II, the differences in the relative difficulty of these forms for their age groups should be pointed out at this time. The ratios recorded in the last column of Table 10B show that the differences between these means were larger than any that have been mentioned previously and were all highly significant. It will be recalled that the tests were so constructed that this combination should have presented the greatest difference between the sets of scores.

IV. THE FACTORIAL ANALYSES

The intercorrelations of the scores of each age group on Form I, on Form II, and on the Combined Form were analyzed by the bi-factor method. These analyses, as are all factor analyses, are somewhat crude, because there is no known way to determine precisely "when to stop factoring." A few general principles are accepted by most workers in this field and the present investigator has been guided by them. The first of these is that factor loadings which are less than .316 generally may be disregarded because they contribute less than 10 per cent to the variance of the test. The second principle is that factor coefficients become questionable when they are less than three times their standard error. This criterion, however, should be used with a great deal of caution because the standard errors may be somewhat too small or the assumptions underlying their use may not always be fulfilled (18, pp. 131-132). The third principle is that the analysis should not be carried so far that insignificant residuals are being analyzed. Again, the test for the significance of a residual is crude (18, p. 131) and should be used with great caution. There is also a subjective element involved in the determination of the significance of a residual because groups of small, but positive, residuals actually may be significant when the individual residuals that

TABLE 11

INTERCORRELATIONS OF SCORES OF BOYS AGED 9 YEARS ON FORM I OF EACH 10 TESTS

Test	1	2	3	4	5	6	7	8	9	10
1										
2	.505									
3	.314	.530								
4	.365	.457	.377							
5	.280	.414	.336	.386						
6	.479	.440	.349	.529	.466					
7	.363	.442	.225	.325	.479	.500				
8	.370	.479	.340	.302	.520	.419	.796			
9	.405	.485	.279	.312	.476	.481	.855	.829		
10	.519	.571	.350	.254	.481	.401	.689	.700	.726	

READ: For each test read along the row to the diagonal and then down in the column carrying the test number. For instance, the correlation between Test 4 and each of the remaining tests from 1 to 10 respectively is read .365, .457, .377, .386, .529, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

TABLE 12

BI-FACTOR PATTERN OF SCORES OF BOYS AGED 9 YEARS ON FORM I OF EACH OF 10 TESTS*

Test	General Factor B_0	Group I Factor B_1	Group II Factor B_2	Group III Factor B_3	Doublet D_1	Unique U_1
1	.613					.790
2	.724					.690
3	.506					.863
4	.544					.839
5	.646					.764
6	.688					.726
7	.620			.651		.438
8	.650			.606		.459
9	.651			.681		.335
10	.684			.414		.601
Contribution of Factor	4.041			1.425		4.534

* $\sigma_a = .088$

READ HORIZONTALLY: For Test 7 the general factor coefficient is .620, the group factor coefficient, .651, and the coefficient of the unique factor .438.

READ VERTICALLY: The coefficient of the general factor for Test 1 is .613; for Test 2, .724, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

form the groups would not be significant if they appeared alone (18, pp. 135-136). Whether "to factor or not to factor" is a decision that must often be made arbitrarily by the person making the analysis. Such decisions have been made by the present investigator in the six solutions presented.

The intercorrelations of the scores of each age group on Form I, Form II, and the Combined Form are given in Tables 11, 13, 15, 17, 19, and 21. The factor patterns derived from each of these tables follows immediately after it and these are numbered Tables 12, 14, 16, 18, 20, and 22, respectively.

A. SUMMARY OF THE SIX BI-FACTOR SOLUTIONS

The factor patterns derived from the six correlation tables showed two types of changes. First, there was a change in the number and kinds of factors which emerged. Second, there was a change in the relative size of the contributions of the general factor, of the group factors, and of their sum, the communality, to the variance of the tests.

TABLE 13

INTERCORRELATIONS OF SCORES OF BOYS AGED 9 YEARS ON FORM II OF EACH OF 10 TESTS

Test	1	2	3	4	5	6	7	8	9	10
1										
2	.374									
3	.239	.291								
4	.265	.453	.024							
5	.286	.529	.186	.397						
6	.229	.388	.019	.450	.515					
7	.125	.238	.244	.263	.445	.215				
8	.238	.379	.237	.310	.401	.380	.634			
9	.256	.526	.102	.272	.505	.312	.530	.625		
10	.423	.531	.092	.375	.560	.365	.435	.554	.681	

READ: For each test read along the row to the diagonal and then down in the column carrying the test number. For instance the correlation between Test 4 and each of the remaining tests from 1 to 10 respectively is read .265, .453, .024, .397, .450, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

TABLE 14

BI-FACTOR PATTERN OF SCORES OF BOYS AGED 9 YEARS ON FORM II OF EACH OF 10 TESTS*

Test	General Factor B_0	Group I Factor B_1	Group II Factor B_2	Group III Factor B_3	Doublet D_1	Unique U_1
1	.430	.251				.867
2	.722	.254				.644
3	.176	.646				.742
4	.517					.856
5	.751					.660
6	.488					.872
7	.485			.474		.734
8	.626			.524		.578
9	.636			.675		.374
10	.756			.222		.616
Contribution of Factor	3.410			1.004		5.040

* $\sigma_a = .113$

READ HORIZONTALLY: For Test 7 the general coefficient is .485, the group factor coefficient, .474, and the coefficient of the unique factor is .734.

READ VERTICALLY: The coefficient of the general factor for Test 1 is .430; for Test 2, .722, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

TABLE 15

INTERCORRELATIONS OF SCORES OF BOYS AGED 9 YEARS ON THE COMBINED FORM OF EACH OF 10 TESTS

Test	1	2	3	4	5	6	7	8	9	10
1										
2	.524									
3	.341	.570								
4	.376	.494.	.318							
5	.352	.545	.386	.433						
6	.428	.469	.301	.555	.558					
7	.311	.436	.245	.334	.552	.421				
8	.368	.502	.359	.336	.552	.444	.837			
9	.402	.570	.339	.343	.561	.459	.844	.844		
10	.572	.628	.390	.350	.605	.428	.698	.738	.770	

READ: For each test read along the row to the diagonal and then down in the column carrying the test number. For instance, the correlation between Test 4 and each of the remaining tests from 1 to 10 respectively is read .376, .494, .318, .433, .555, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

Table 23 shows all of the factor patterns which emerged in this experiment. Form I was factorially simple since at both age levels only the general factor and one group factor were required to describe its communality. Form II was more complex. At nine years of age the general factor and two group factors plus a doublet were required. At both age levels the Combined Form necessitated the use of the general factor and one group factor plus a doublet to describe its communality.

It will be noted that on the 12-year-level this doublet appeared in the group where a factor had emerged on Form II for 9-year-olds, and that on the 9-year-level it appeared in the group where a factor had emerged on Form II for the 12-year-olds. The appearance of a factor in one case and a doublet in the other was due mathematically to fluctuations in the relative sizes of group residuals. Whether these fluctuations were due to chance or whether they had psychological significance can be ascertained only by further research.

The communalities of the forms for both age groups and the contributions of the general and group factors to the variance of each are shown in Table 24. The inclusion of the large doublet in the communality of the Combined Form for 12-year-olds makes the quantitative interpretation of this

TABLE 16

BI-FACTOR PATTERN OF SCORES OF BOYS AGED 9 YEARS ON COMBINED FORM OF EACH OF 10 TESTS*

Test	General Factor B_0	Group I Factor B_1	Group II Factor B_2	Group III Factor B_3	Doublet** D_1	Unique U_1
1	.596					.803
2	.792					.611
3	.499					.867
4	.562				.324	.827
5	.754					.657
6	.640				.324	.768
7	.596			.685		.421
8	.668			.622		.409
9	.698			.626		.348
10	.774			.356		.523
Contribution of Factor	4.411			1.376	.210	4.213

* $\overset{\sigma}{a} = .079$

** This doublet was established at a low confidence level and is not included in further analyses.

READ HORIZONTALLY: For Test 7 the general factor coefficient is .596, the group factor coefficient, .685, and the coefficient of the unique factor, .421.

READ VERTICALLY: The coefficient of the general factor for Test 1 is .596; for Test 2, .792, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

battery both difficult and uncertain. Therefore, the Combined Form for both nine and 12-year-old boys was not subjected to further analysis.

The data in Table 24 may be examined to determine what changes in the contributions of the general and of the sums of the group factors occurred with change in age. There were changes between the age of nine and 12 years in each of the two forms. In Form I the general factor increased 0.484 or approximately 11.2 per cent and in Form II, the general factor increased 0.533 or 15.6 per cent. The sums of the group factors decreased with age in both forms. In Form I the decrease amounted to 0.785 or 55.1 per cent. In Form II the decrease was 0.356 or approximately 23 per cent. The communality, which is the sum of the general and group factors, showed a small decrease with age in Form I and a small increase with age in Form II. The

TABLE 17

INTERCORRELATIONS OF SCORES OF BOYS AGED 12 YEARS ON FORM I OF EACH OF 10 TESTS

Test	1	2	3	4	5	6	7	8	9	10
1										
2	.494									
3	.302	.400								
4	.486	.467	.263							
5	.405	.310	.236	.470						
6	.410	.427	.203	.392	.369					
7	.456	.500	.347	.506	.508	.250				
8	.535	.514	.411	.561	.450	.314	.719			
9	.527	.480	.313	.522	.442	.257	.744	.699		
10	.643	.491	.384	.554	.523	.426	.658	.725	.670	

READ: For each test read along the row to the diagonal and then down in the column carrying the test number. For instance, the correlation betwen Test 4 and each of the remaining tests from 1 to 10 respectively is read .486, .467, .263, .470, .392, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

decrease amounted to 0.301 or 5.5 per cent in Form I and the increase to 0.1777 or 3.4 per cent in Form II. In these analyses, the content of the tests was the same for each age group, but both age and relative difficulty were permitted to vary. It is, however, the method used in all factor studies reported in this study, and the results are not in conformity with those to be expected by the proposed differentiation hypothesis.

The design of this experiment permits us to study the shifts in the contributions of general and of the sums of the group factors under conditions in which age is permitted to vary but in which differences in relative difficulty is minimized, and in which the content for the two populations is similar, but not identical.

The effects of increasing age upon the factor patterns under those conditions can be shown by comparing the factor pattern of boys aged nine years on Form I with the pattern of boys aged 12 years on Form II. This comparison shows that under conditions of minimum difference in difficulty the contribution of the general factor to the variance of the tests was 0.098 less for the older than for the younger boys and, at the same time, the contribution of the sum of the group factors was 0.231 less for the older boys. Using the contribution of the general factor at age nine as the base,

TABLE 18

BI-FACTOR PATTERN OF SCORES OF BOYS AGED 12 YEARS ON FORM I OF EACH OF 10 TESTS*

Test	General Factor B_0	Group I Factor B_1	Group II Factor B_2	Group III Factor B_3	Doublet D_1	Unique U_1
1	.727					.689
2	.671					.742
3	.439					.899
4	.682					.731
5	.571					.820
6	.481					.876
7	.717			.431		.549
8	.777			.389		.495
9	.709			.480		.517
10	.844			.270		.465
Contribution of Factor	4.525			0.640		4.835

* $a = .082$

READ HORIZONTALLY: For Test 7 the general factor coefficient is .717, the group factor is .431, and the coefficient of the unique factor is .549.

READ VERTICALLY: The coefficients of the general factor for Test 1 is .727; for Test 2, .671, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

we may note that the decrease of 0.098 amounted to 2.4 per cent. Using the contribution of the sum of the group factors at age nine as the base, we note that the decrease of 0.231 amounted to 16.2 per cent. It should be recalled that the contents of the two forms were similar, but not identical. It should also be emphasized that there is no known test for the significance of these differences, even if the content were identical.

These results suggest that age alone may make little difference in the size of the general factor. It may influence the group factors to a greater extent.

The data in Table 24 may be examined to determine the effect of differences in test difficulty upon the factor patterns. This can be shown by comparing the patterns of each age group on each of the two forms. Based upon the contribution of the general factor in Form I at the respective age level as the unit of measure, the general factor for boys aged nine years decreased 15.6 per cent with the increase in difficulty between Form I and Form II,

TABLE 19

INTERCORRELATIONS OF SCORES OF BOYS AGED 12 YEARS ON FORM II OF EACT OF 10 TESTS

Test	1	2	3	4	5	6	7	8	9	10
1										
2	.528									
3	.288	.514								
4	.438	.527	.284							
5	.331	.271	.096	.467						
6	.371	.450	.255	.422	.357					
7	.330	.386	.249	.303	.181	.234				
8	.394	.428	.177	.446	.385	.304	.461			
9	.442	.516	.401	.510	.383	.375	.550	.511		
10	.526	.622	.420	.503	.433	.363	.554	.555	.623	

READ: For each test read along the row to the diagonal and then down in the column carrying the test number. For instance, the correlation between Test 4 and each of the remaining tests from 1 to 10 respectively is read .438, .527, .284, .467, .422, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

and the general factor for boys aged 12 years decreased 12.9 per cent under the same conditions. Based upon the contribution of the sum of the group factors in Form I at the respective age level as the unit of measure, the contribution of the group factors for boys aged nine years increased 8.8 per cent and the contribution of the group factors for boys aged 12 years increased 87 per cent under the same conditions. Again it should be pointed out that a portion of the differences noted may be due to the fact that the two forms were similar in content but were not identical. They certainly differed in difficulty.

When age is permitted to vary and the extremes of relative difficulty are analyzed, the trends noted above are even more pronounced. The greatest differences in relative difficulty are those between the 12-year-olds on Form I and 9-year-olds on Form II. Using the contribution of the general factor at age 12 years on Form I as a unit of measure, we may see that the increase in relative difficulty was accompanied by a decrease of 24.6 per cent in the contribution of the general factor. Using the contribution of the sum of the group factors to the variance of the same tests at age 12 years as a unit of measure, we note that the increase in difficulty was accompanied by an increase of 142 per cent in the contribution of the group factors.

TABLE 20

BI-FACTOR PATTERN OF SCORES OF BOYS AGED 12 YEARS ON FORM II OF EACH OF 10 TESTS*

Test	General Factor B_0	Group I Factor B_1	Group II Factor B_2	Group III Factor B_3	Doublet D_1	Unique U_1
1	.661					.751
2	.742					.670
3	.423					.906
4	.671		.302			.677
5	.444		.559			.700
6	.530		.218			.820
7	.479			.803	.214**	.355
8	.612			.209		.762
9	.755			.234		.612
10	.821				.214**	.571
Contribution of Factor	3.943		.451	.743	.092	4.863

* $\sigma_a = .095$

** Due to its size, this doublet is not included in the communality reported for these tests.

READ HORIZONTALLY: For Test 7 the general factor coefficient is .479, the group factor coefficient is .803, and the coefficient of the unique factor is .355. The doublet of .214 is not included in the communality.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

TABLE 21

INTERCORRELATIONS OF SCORES OF BOYS AGED 12 YEARS ON COMBINED FORM OF EACH OF 10 TESTS

Test	1	2	3	4	5	6	7	8	9	10
1										
2	.570									
3	.319	.759								
4	.514	.557	.325							
5	.406	.347	.193	.508						
6	.427	.470	.282	.460	.412					
7	.452	.505	.352	.466	.430	.284				
8	.533	.548	.367	.573	.461	.352	.713			
9	.539	.569	.393	.559	.472	.374	.720	.677		
10	.670	.609	.465	.571	.529	.445	.683	.727	.699	

READ: For each test read along the row to the diagonal and then down in the column carrying the test number. For instance, the correlation between Test 4 and each of the remaining tests from 1 to 10 respectively is read .514, .557, .325, .508, .460, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

TABLE 22

BI-FACTOR PATTERN OF SCORES OF BOYS AGED 12 YEARS ON COMBINED FORM OF EACH OF 10 TESTS*

Test	General Factor B_0	Group I Factor B_1	Group II Factor B_2	Group III Factor B_3	Doublet D_1	Unique U_1
1	.730					.684
2	.743				.550**	.382
3	.478				.550**	.685
4	.710		.282			.645
5	.546		.424			.722
6	.535		.282			.796
7	.662			.690		.292
8	.755			.309		.579
9	.774			.301		.557
10	.875					.485
Contribution of Factor	4:775		.340	.662	.605	3.638

* $\sigma_a = .074$

** This doublet is included in the communality reported for these tests.

READ HORIZONTALLY: For Test 7 the general factor coefficient is .662, the group factor coefficient, .690, and the coefficient of the unique factor .292.

READ VERTICALLY: The coefficient of the general factor for Test 1 is .730; for Test 2, .743, etc.

NAMES OF TESTS

1. Finding Middle Number
2. Number Series
3. Number Patterns
4. Figure Exercises
5. Figure Dividing
6. Boys and Girls
7. Vocabulary
8. Opposites
9. Word Grouping
10. Arithmetic Reasoning

TABLE 23

SUMMARY OF FACTOR PATTERNS DERIVED FROM THE SIX CORRELATION TABLES OF THIS EXPERIMENT

Age	Form	General	Group I	Group II	Group III
9	I	X			X
12	I	X			X
9	II	X	X		X
12	II	X		X	X, D*
9	C	X		D*	X
12	C	X	D**		X

* This doublet was established at a low confidence level and was not included in further analyses.

** This doublet was established at a high confidence level and was included in further analyses.

READ: For subjects aged nine years, Form I showed a general and a Group III factor. . . . For subjects aged 12 years, Form II showed a general factor, a Group II factor, and a factor plus a doublet in Group III.

These results suggest that increasing the relative difficulty of tests tends to decrease the contribution of the general factor to the variance of the tests and, at the same time, to increase the contribution of the group factors.

TABLE 24

CONTRIBUTIONS TO THE TEST VARIANCES OF THE GENERAL FACTORS, OF THE SUMS OF THE GROUP FACTORS AND OF THE COMMUNALITIES FOR BOYS AGED 9 YEARS AND FOR BOYS AGED 12 YEARS ON EACH FORM

Age	Form	Contribution to Test Variance		
		General Factor	Sums of Group Factors	Communality
9	I	4.041	1.425	5.466
12	I	4.525	.640	5.165
9	II	3.410	1.550	4.960
12	II	3.943	1.194	5.137*
9	C	4.411	1.376	5.787*
12	C	4.775	1.607	6.382**

* Doublet excluded from the communality.
** Doublet included in the communality.

READ HORIZONTALLY: For subjects aged nine years on Form I the general factor contributed 4.041, the sum of the group factors contributed 1.425, and the total of the two, the communality, contribtued 5.466 to the variances of the tests.

READ VERTICALLY: For subjects aged nine years on Form I the general factor contributed 4.041; for subjects aged 12 years on Form I the general factor contributed 4.525 to the variance of the tests; etc.

V. INTERPRETATIONS AND CONCLUSIONS

A. The First Hypothesis

In this chapter we are primarily concerned with the interpretation of these data in relation to the hypotheses stated in Section I. It will be recalled that the first hypothesis, that proposed by Professor Garrett, was: "Abstract or symbol intelligence changes in its organization as age increases from a fairly unified and general ability to a loosely organized group of abilities or factors" (12, p. 373). By this hypothesis, as the age of the subjects increased the general factor should have contributed less to the variance of the tests. Between the ages of nine and 12 years in both Form I and Form II there was, on the contrary, an increase in the contribution of the general factor. By the same hypothesis, there should have been an increase with increase in age in the contribution of the group factors to the variance of the tests. In this study there was a decrease with increase in age in the contribution of such factors. Furthermore, an increase in the specificity of the tests and an accompanying decrease in their communalities would have supported this hypothesis. The communalities, however, remained relatively stable, those in Form I showing slight decrease with age and those in Form II showing slight increase with age. It is apparent that those data do not support the hypothesis stated above.

If these data are confirmed by other studies it would appear that the first hypothesis will need to be re-examined.

B. The Second Hypothesis

The second hypothesis, that proposed by the investigator, was: The observed differences between the factor patterns for two age groups are due not only to the age differences of the groups in question but also to differences in difficulty levels of the test materials for the two populations. Since the design of this experiment did not make it possible to hold both difficulty and content constant and to permit age, alone, to vary, we are unable to state that age, alone, did or did not affect the observed shifts in the factor patterns. When the difference in relative difficulty was minimized, the general factor was quite stable, showing a decrease of only 2.4 per cent between the ages of nine and 12, but the sums of the contributions of group factors showed a decrease of 16.2 per cent. It may be that group factors are more sensitive to changes that occur with normal increase in age than is the general factor.

There is strong evidence to support the hypothesis that the observed differ-

ences between the factor patterns for two age groups are due in part to differences in the relative difficulty of the materials. The analysis in the concluding section of Chapter IV showed that every time a test situation became relatively easier there was an increase in the contribution of the general factor and a decrease in the contribution of the sum of the group factors. It is certain that consistent changes in the factor patterns derived from those data accompanied known changes in the difficulty levels of these materials for the subjects of this experiment.

C. Implications for Further Research

Since these data do not conform with the data obtained by other investigators, the scores of the girls who took these tests should be analyzed to determine whether sex differences may have been the disturbing influence.

The influences that control the communality of tests should also be subjected to further study. In this study the communalities obtained showed small and contradictory changes with age. The communality of the easier form was somewhat greater for subjects of both ages than was the communality of the more difficult form. In this study, the communalities were relatively stable, being influenced very little by either changes in age or in the relative difficulty of the test materials. These data are too meager to warrant sweeping generalizations, however.

The influence of age and of test difficulty upon the size and the number of the group factors is also noteworthy. In every case the contribution of the group factors was greater at the age of nine years than at the age of 12 years. The contribution of the group factors was greater in Form II than in Form I at each age level. There were also more group factors in Form II than in Form I at each age level. It thus appears that relative immaturity on the part of the subjects and greater difficulty on the part of the materials may call into play quite highly specialized abilities which, among more mature subjects and with easier materials, become a part of the general competence. This may result in part from physiological maturation and in part from a longer common school training and the increase in experience which normally accompanies an increase in age in a typical American community.

Closely related to the appearance or disappearance of group factors is the change in the specific tests which make up the group in which the factors do appear. Other investigators have found that tests commonly shift from one group to another depending upon the characteristics of the population tested (3). It would be extremely difficult to determine with any degree of certainty when such changes are part of a directed process and when they are the

results of chance. However these shifts within the common factor space indicate that the factor pattern is not a static structure. The study of the dynamics of the factor pattern thereby becomes an important problem for further research.

In conclusion it should be stated that by using different methods of factor analysis of data such as these and by comparing the results obtained, one might contribute to the understanding of the changes which occur in the factor pattern with changes in age and in test difficulty.

REFERENCES

1. Asch, S. E. A study of change in mental organization. *Arch. of Psychol.,* 1936, **28**, 17-25.
2. Babcock, H., & Emerson, M. An analytical study of the MacQuarrie test for mechanical ability. *J. Educ. Psychol.,* 1938, **29**, 50-55.
3. Balinsky, B. An analysis of the mental factors of various age groups from nine to sixty. *Genet. Psychol. Monog.,* 1941, **23**, 191-234.
4. Blumenfeld, W. The invariability of certain coefficients of correlation during human development. *J. Genet. Psychol.,* 1946, **68**, 189-204.
5. Bryan, A. I. Organization of memory in younger children. *Arch. of Psychol.,* 1934, **24**, No. 162.
6. Burt, C., & John, E. A factorial analysis of Terman Binet tests: Part II. *Brit. J. Educ. Psychol.,* 1942, **12**, 156-161.
7. Clark, M. P. Changes in primary mental ability with age. *Arch. of Psychol.,* 1944, **41**, No 291.
8. Cook, W. W. "The measurement of general spelling ability involving controlled comparisons between techniques." Iowa City: *Univ. Iowa Stud. of Educ.,* 1932, **6**, No. 6.
9. Dunlap, J. W., & Kurtz, A. K. Handbook of Statistical Nomographs, Tables, and Formulas. Yonkers-on-Hudson, New York: World Book, 1932.
10. Ferguson, G. A. The factorial interpretation of test difficulty. *Psychometrika,* 1941, **6**, 323-329.
11. Garrett, H. E. Differentiable mental traits. *Psychol. Rec.,* 1938, **2**, 259-298.
12. ————. A developmental theory of intelligence. *Amer. Psychol.,* 1946, **1**, 372-378.
13. Garrett, H. E., Bryan, A. I., & Perl, R. E. The age factor in mental organization. *Arch. of Psychol.,* 1935, **26**, No. 176.
14. Goodenough, F. L., & Maurer, K. M. The mental growth of children from two to fourteen years: A study of the predictive value of the Minnesota pre-school scales. *Univ. Minn. Inst. Child Welf. Monog.,* 1942, No. 19.
15. Guilford, J. P. The difficulty of a test and its factor composition. *Psychometrika,* 1941, **6**, 67-77.
16. Guttman, L. A basis for analyzing test-retest reliability. *Psychometrika,* 1945, **10**, 255-282.
17. Hertzman, M. The effects of the relation difficulty of mental tests on patterns of mental organization. *Arch. of Psychol.,* 1936, **28**, No. 197.
18. Holzinger, K. J., & Harmon, H. H. Factor Analysis. Chicago: Univ. Chicago Press, 1941.
19. Kelley, T. L. Crossroads in the Mind of Man. Stanford University, California: Stanford Univ. Press, 1928.
20. ————. Comment on Wilson and Worcester's "Note on factor analysis." *Psychometrika,* 1940, **5**, 117-120.
21. Kenney, J. F. Mathematics of statistics. New York: D. Van Nostrand, 1939.
22. McNemar, Q. The Revision of the Stanford-Binet Scale: An analysis of the standardization data. Boston: Houghton Mifflin, 1942.
23. Morrow, R. S. An experimental analysis of the theory of independent abilities. *J. Educ. Psychol.,* 1941, **32**, 495-512.
24. Odell, C. W. Statistical Methods in Education. New York: D. Appleton-Century, 1935.
25. Raven, J. C. The R.E.C.I. series of perceptual tests: An experimental survey. *Brit. J. Med. Psychol.,* 1939, **18**, 16-34.

26. REICHARD, S. Mental organization and age level. *Arch. of Psychol.,* 1944, **41**, No. 295.

27. RICHARDS, T. W. Genetic emergence of factor specificity. *Psychometrika,* 1941, **6**, 37-42.

28. SARGENT, S. S. Effects of difficulty level upon thinking processes. *Arch. of Psychol.,* 1940, **35**, No. 249.

29. SCHILLER, B. Verbal, numerical, and spatial abilities of young children. *Arch. of Psychol.,* 1934, **24**, No. 161.

30. SCHNECK, M. M. R. The measurement of verbal and numberical abilities. *Arch. of Psychol.,* 1929, **17**, No. 107.

31. SLOCOMBE, C. S. The constancy of "*g*," general intelligence. *Brit. J. Psychol.,* 1926, **17**, 93-110.

32. SMITH, G. M. Group factors in mental tests similar in material or in structure. *Arch. of Psychol.,* 1933, **24**, No. 159.

33. SWINEFORD, F. Growth in the general and verbal bi-factors from Grade VII to Grade IX. *J. Educ. Psychol.,* 1947, **38**, 257-272.

34. SWINEFORD, F., & HOLZINGER, K. J. A study in factor analysis: The reliability of bi-factors and their relation to other measures. *Suppl. Educ. Monog.,* 1942, No. 53.

35. THURSTONE, L. L. Primary mental abilities. *Psychomet. Monog.,* 1938, No. 1.

36. WRIGHT, R. A factor analysis of the original Stanford-Binet scale. *Psychometrika,* 1939, **4**, 209-220.

VITA

Hazen Alonzo Curtis was born September 18, 1905. He attended Evansville College and Southern Illinois Normal University. The Bachelor of Education degree was conferred by the latter institution in 1934. The degree of Master of Science in Education was conferred by the University of Illinois in 1938, and the Master of Education degree by the same institution in 1947. He is a member of Kappa Phi Kappa, Kappa Delta Pi, and Phi Delta Kappa. He is the author of "Wide Reading for Beginners," *Journal of Educational Research,* 1938, 32, 252-262; and, with E. W. Dolch, of "Do Spelling Books Teach Spelling?" *Elementary School Journal,* 1939, 39, 584-592. He taught in the rural schools of Edwards County, Illinois, from 1925-1927; served as Principal of an elementary school in Staunton, Illinois, from 1927-1934, and as Superintendent of the same schools from 1934-1942. He was Superintendent of Schools at Knoxville, Illinois, during the year 1942-1943. He was on active duty in the Naval Reserve from 1943-1946, holding the rank of Lieutenant, USNR. He was Professor of Education and Head of the Department in Arizona State College, Flagstaff, during the academic year 1947-1948, and now is Professor of Education in Florida State University, Tallahassee.